

# Reconocimiento automático del alfabeto doctilológico.

Chyzhyk Darya

December 23, 2008

## Abstract

En este papel se presenta un estudio sobre la construcción de un clasificador automático para reconocer visualmente los signos del lenguaje de sordos correspondientes a las letras del alfabeto. Se muestra una metodología de aplicación del análisis de correlación a un proceso de reconocimiento visual de signos.

## Introducción

La tarea consiste en reconocer el lenguaje visual de las personas sordas. En el mundo hay mucha gente sordomuda (con defectos del lenguaje). El desarrollo de la ciencia y de técnica crea las condiciones para invención de nuevos medios informáticos de comunicación e interacción hombre-máquina. La forma habitual de comunicación de las personas sordas es el lenguaje de los gestos. Por lo tanto mi tarea se reduce a que escribir un programa para reconocer automáticamente los gestos del lenguaje de los sordos.

Tenemos una base de datos experimental constituida por las fotografías obtenidas con una persona voluntaria de la Asociación de Personas Sordas de Gipuzkoa mostrando los signos correspondientes a las letras del abecedario. Consta de cinco carpetas. En cada carpeta hay 25 fotos de la chica voluntaria, que muestra las 25 letras: a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p, q, r, s, t, u, w, x, y1, y2. Las carpetas se distinguen por modo de mostrar los gestos. La primera carpeta usamos para la formación de un patrón, las restantes carpetas se emplean como test para el reconocimiento basado en dicho patrón.

En el fondo, mi tarea se consiste en producir y ejecutar dos programas consecutivos. El primero encuentra la mano de la persona entre otros objetos en la foto. El segundo programa sirve para clasificar el gesto de la mano extraída mediante la comparación con la base de datos de patrones.

# Extracción de la mano

He hecho un programa, que se llama "Selección mano fajar". Este función selecciona el objeto "mano" entre los otros objetos en la imagen, lo extrae con su mínimo rectángulo envolvente. Cada letra tiene su rectángulo envolvente con un tamaño distinto que depende de tamaño de la mano en la imagen. Para utilizar el análisis de correlación es necesario tener las imágenes del mismo tamaño. Otra parte de esta función mete la mano extraída con sus parámetros de tamaño en el centro de un rectángulo negro con tamaño 180x160.

Por ejemplo, tenemos la imagen de la letra "a":



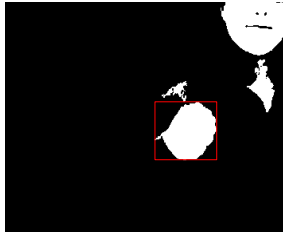
Transformamos la imagen en imagen binaria.



Removemos la basura digital (es decir, los puntos pequeños blancos que no pertenecen a la imagen de la mano).



Identificamos la mano entre otros objetos grandes como mas cercana al centro del imagen y seleccionamos un rectángulo mínimo en que podemos meter la mano.



Cogemos la mano.



Y la metemos en el centro del rectángulo negro.



Lo mismo hacemos con las letras b, c, d, r, s, t, u:  
La letra "b"



La letra "c"



La letra "d"



La letra "r"



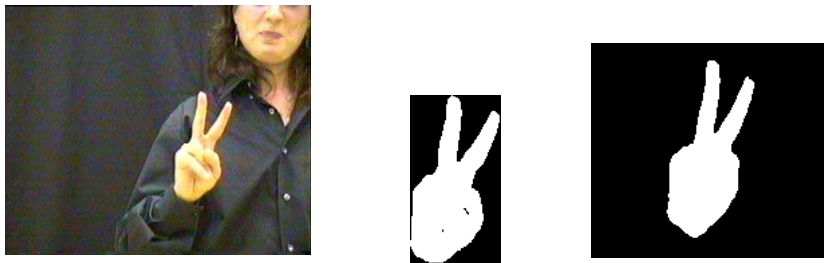
La letra "s"



La letra "t"



La letra "u"



Mediante este programa tratando las fotos de primera carpeta formamos la base de datos de los patrones.



letra a



letra b



letra c



letra d



letra e



letra f



letra g



letra h



letra i



letra j



letra k



letra l



letra n



letra n



letra o



letra p



letra q



letra r



letra s



letra t



letra u



letra w



letra x



letra y1



letra y2

## Experimento de clasificación.

Además he escrito un programa que se llama “identificar mano”. Este programa compara la mano extraída mediante el programa “Selección mano fija” de una imagen de prueba con todos los patrones extraídos de la primera carpeta de la base de datos (25) y nos da como respuesta la letra a la que corresponde esta imagen.

El proceso de la comparación se realiza usando las formulas de correlación.

Suponemos que  $X(x_1, \dots, x_n)$  - es la imagen que queremos clasificar,  $Y(y_1, \dots, y_n)$  - es la imagen del patrón. El coeficiente de correlación entre estos vectores esta dado por la fórmula siguiente:

$$\text{corr}(X, Y) = \frac{\text{cov}(X, Y)}{\sigma_x \sigma_Y} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}$$

Si  $\text{corr}(X, Y) = 1$ , ambos vectores son colineales (paralelos).

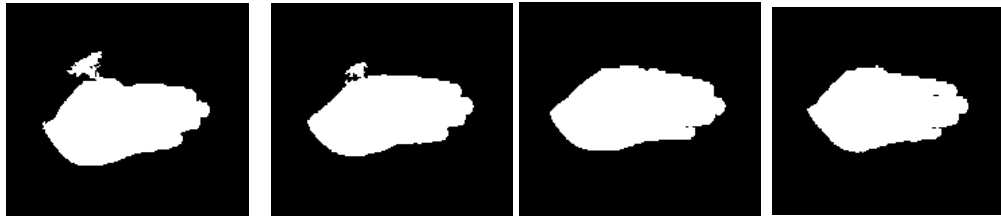
Si  $\text{corr}(X, Y) = 0$ , ambos vectores son ortogonales.

Si  $\text{corr}(X, Y) = -1$ , ambos vectores son colineales de dirección opuesta.

Por lo tanto podemos decir, que las imágenes que tienen el coeficiente de correlación lo más alto, son más parecidos que otros con menor coeficiente de correlación.

En las siguientes imágenes se presentan los ejemplos de los resultados: las imágenes de la misma letra pero de diferentes realizaciones con coeficientes de correlación de cada letra del patrón. En negrita mostramos los máximos que son las respuestas del programa, identificando la letra que ha sido reconocida. En las siguientes tablas cada fila corresponde a una realización de la letra extraída de una carpeta distinta. Las columnas corresponden a cada patrón extraído de la primera carpeta.. Se muestran los resultados para cada una de las letras.

La letra “b”



Letra	a	<b>b</b>	c	d	e	f	g	h
corr b2	0.5651	<b>0.7439</b>	0.5270	0.5557	0.5882	0.5137	0.6341	0.6411
corr b3	0.6081	<b>0.8675</b>	0.4918	0.4867	0.5573	0.4702	0.6900	0.6269
corr b4	0.5896	<b>0.9031</b>	0.4545	0.4217	0.5147	0.4448	0.6612	0.5643
corr b5	0.6054	<b>0.9460</b>	0.4722	0.4417	0.5408	0.4782	0.6781	0.5835

Letra	i	j	k	l	m	n	o	p
corr b2	0.6035	0.4981	0.6104	0.4485	0.3258	0.3203	0.5266	0.3571
corr b3	0.5521	0.5136	0.7164	0.3584	0.3803	0.3885	0.5963	0.3178
corr b4	0.4877	0.5088	0.7306	0.2978	0.3983	0.4172	0.5931	0.2775
corr b5	0.5256	0.5321	0.7345	0.3129	0.4065	0.4229	0.6047	0.2959

Letra	q	r	s	t	u	w	x	y1	y2
corr b2	0.4628	0.3873	0.5853	0.5364	0.3902	0.3189	0.4704	0.5840	0.5898
corr b3	0.4732	0.3604	0.5358	0.4844	0.3619	0.3082	0.4795	0.5647	0.5875
corr b4	0.4516	0.3304	0.5083	0.4571	0.3287	0.2959	0.4716	0.5253	0.5694
corr b5	0.4666	0.3463	0.5469	0.4971	0.3488	0.3171	0.4913	0.5518	0.5939

La respuesta del programa en todos los casos es “letra identificada = b”

Letra “c”



Letra	a	b	<b>c</b>	d	e	f	g	h
corr c2	0.5165	0.5132	<b>0.8627</b>	0.5610	0.7999	0.5818	0.6235	0.6455
corr c3	0.5732	0.4663	<b>0.8836</b>	0.5756	0.7438	0.5027	0.6159	0.6912
corr c4	0.5006	0.4921	<b>0.7841</b>	0.6345	0.7626	0.5721	0.6243	0.6586
corr c5	0.5768	0.4453	<b>0.8754</b>	0.5812	0.7391	0.4932	0.5767	0.6810

Letra	i	j	k	l	m	n	o	p
corr c2	0.5342	0.5896	0.5125	0.5318	0.3203	0.3074	0.6654	0.3620
corr c3	0.5147	0.6036	0.5631	0.5893	0.3270	0.2996	0.6748	0.4156
corr c4	0.4969	0.5805	0.4967	0.5449	0.2936	0.2742	0.6909	0.3556
corr c5	0.5294	0.6013	0.5296	0.5680	0.2996	0.2673	0.6602	0.4227



Letra	q	r	s	t	u	w	x	y1	y2
corr c2	0.5811	0.4477	0.5637	0.4998	0.4755	0.4379	0.5297	0.5460	0.6301
corr c3	0.6196	0.4857	0.4810	0.4191	0.5139	0.4471	0.5188	0.5385	0.5986
corr c4	0.5776	0.4442	0.5696	0.4597	0.4934	0.4196	0.4861	0.4995	0.5726
corr c5	0.6174	0.4874	0.4619	0.4255	0.5243	0.4501	0.5222	0.5532	0.6031

La respuesta del programa es “letra identificada = c”

La letra “d”



Letra	a	b	c	d	e	f	g	h
corr d2	0.5576	0.3934	0.4554	0.6403	0.4276	0.5379	0.4311	0.6256
corr d3	0.5787	0.4219	0.5111	<b>0.8101</b>	0.5467	0.5987	0.5013	0.6771
corr d4	0.5379	0.4375	0.4922	<b>0.8409</b>	0.5523	0.6144	0.5345	0.6495
ccor d5	0.5090	0.3960	0.5097	<b>0.8761</b>	0.5601	0.6076	0.4879	0.6325

Letra	i	j	k	l	m	n	o	p
corr d2	0.6443	0.5616	0.3445	0.5260	0.3084	0.2785	0.4708	0.7020
corr d3	0.6617	0.5756	0.3430	0.6714	0.2659	0.2602	0.5159	0.5889
corr d4	0.6171	0.5373	0.3672	0.6787	0.2583	0.2614	0.5149	0.4873
ccor d5	0.6044	0.5197	0.3204	0.6736	0.2310	0.2400	0.4976	0.4815

Letra	q	r	s	t	u	w	x	y1	y2
corr d2	0.6224	<b>0.7226</b>	0.4673	0.4669	0.6008	0.5522	0.5296	0.5955	0.5216
corr d3	0.6425	0.6787	0.5370	0.5540	0.6561	0.5062	0.5769	0.6177	0.5455
corr d4	0.5799	0.5750	0.5558	0.5602	0.6387	0.4780	0.5419	0.5712	0.5258
ccor d5	0.5705	0.5808	0.5319	0.5299	0.6309	0.4674	0.5211	0.5516	0.5040

La respuesta del programa en la experiencia con letra “d2” es “letra identificada = r” en todos otros “d3”-“d5” los casos es “letra identificada = d”.

La letra “s”



Letra	a	b	c	d	e	f	g	h
corr s2	0.4686	0.5659	0.5574	0.6057	0.6768	0.7166	0.5718	0.5704
corr s3	0.4612	0.5849	0.5080	0.5830	0.6504	0.6973	0.5732	0.5678
corr s4	0.4327	0.5611	0.4866	0.5538	0.6283	0.7010	0.5686	0.5521
corr s5	0.4472	0.5582	0.4574	0.5327	0.6065	0.6993	0.5834	0.5089

Letra	i	j	k	l	m	n	o	p
corr s2	0.6016	0.5550	0.4584	0.5462	0.3577	0.3555	0.5481	0.3907
corr s3	0.5996	0.5325	0.4687	0.5130	0.3385	0.3350	0.5412	0.3785
corr s4	0.5691	0.5054	0.4606	0.5029	0.3212	0.3181	0.5366	0.3354
corr s5	0.5906	0.5042	0.4861	0.4981	0.2934	0.2860	0.5050	0.3366

Letra	q	r	s	t	u	w	x	y1	y2
corr s2	0.5365	0.4543	<b>0.8563</b>	0.6951	0.5086	0.4774	0.5896	0.5894	0.6598
corr s3	0.5125	0.4517	<b>0.8770</b>	0.6855	0.4778	0.4673	0.5816	0.5760	0.6357
corr s4	0.4753	0.4076	<b>0.9353</b>	0.6891	0.4542	0.4426	0.5503	0.5678	0.6236
corr s5	0.4593	0.4246	<b>0.8664</b>	0.7447	0.4725	0.4565	0.5822	0.5849	0.6432

La respuesta del programa en todos los casos es “letra identificada = s”

La letra “t”



Letra	a	b	c	d	e	f	g	h
corr t2	0.5622	0.5338	0.4918	0.5051	0.5859	0.7071	0.5944	0.4805
corr t3	0.4977	0.5129	0.4520	0.5104	0.5836	0.6930	0.5795	0.4504
corr t4	0.4770	0.5066	0.4364	0.4968	0.5740	0.6992	0.5772	0.4505
corr r5	0.4832	0.4764	0.4226	0.4904	0.5480	0.7025	0.5458	0.4490

Letra	i	j	k	l	m	n	o	p
corr t2	0.6962	0.6622	0.5069	0.4758	0.3712	0.3684	0.5566	0.4747
corr t3	0.6896	0.5922	0.4479	0.5070	0.3192	0.3037	0.4987	0.4273
corr t4	0.6560	0.5557	0.4571	0.5084	0.3019	0.2965	0.4953	0.3874
corr r5	0.7047	0.5789	0.4262	0.4901	0.2997	0.2815	0.4659	0.4482

Letra	q	r	s	t	u	w	x	y1	y2
corr t2	0.6493	0.5906	0.7045	<b>0.8388</b>	0.6317	0.6377	0.7127	0.6916	0.7353
corr t3	0.5532	0.5114	0.7039	<b>0.9328</b>	0.6008	0.5711	0.6563	0.6407	0.6770
corr t4	0.5082	0.4528	0.7212	<b>0.9367</b>	0.5564	0.5296	0.6212	0.6216	0.6736
corr r5	0.5304	0.4927	0.6675	<b>0.9066</b>	0.6021	0.5630	0.6449	0.6613	0.6771

La respuesta del programa en todos los casos es “letra identificada = t”

Podemos ver que el clasificador construido funciona de modo aceptable. En unos casos el nos da los resultados incorrectos. Sólo en algunos casos aislados, el programa reconoce incorrectamente la letra correspondiente.

## Conclusiones

Este trabajo muestra la descripción del algoritmo que utilizo y los resultados obtenidos. Entendiendo que el objetivo de este trabajo era construir el clasificador de las imágenes de los signos, se ha comenzado por construir el programa que extrae la mano de la imagen original. Es una etapa muy importante y robusta. Utilizando este programa formamos el patrón del alfabeto dactilológico. La siguiente parte del programa establece la correspondencia de una imagen de la mano a una letra del alfabeto. El programa funciona bien pero con algunas equivocaciones. Por eso podemos decir que solo utilizar la correlación no es suficiente para alcanzar un resultado excelente de un 100% de aciertos.